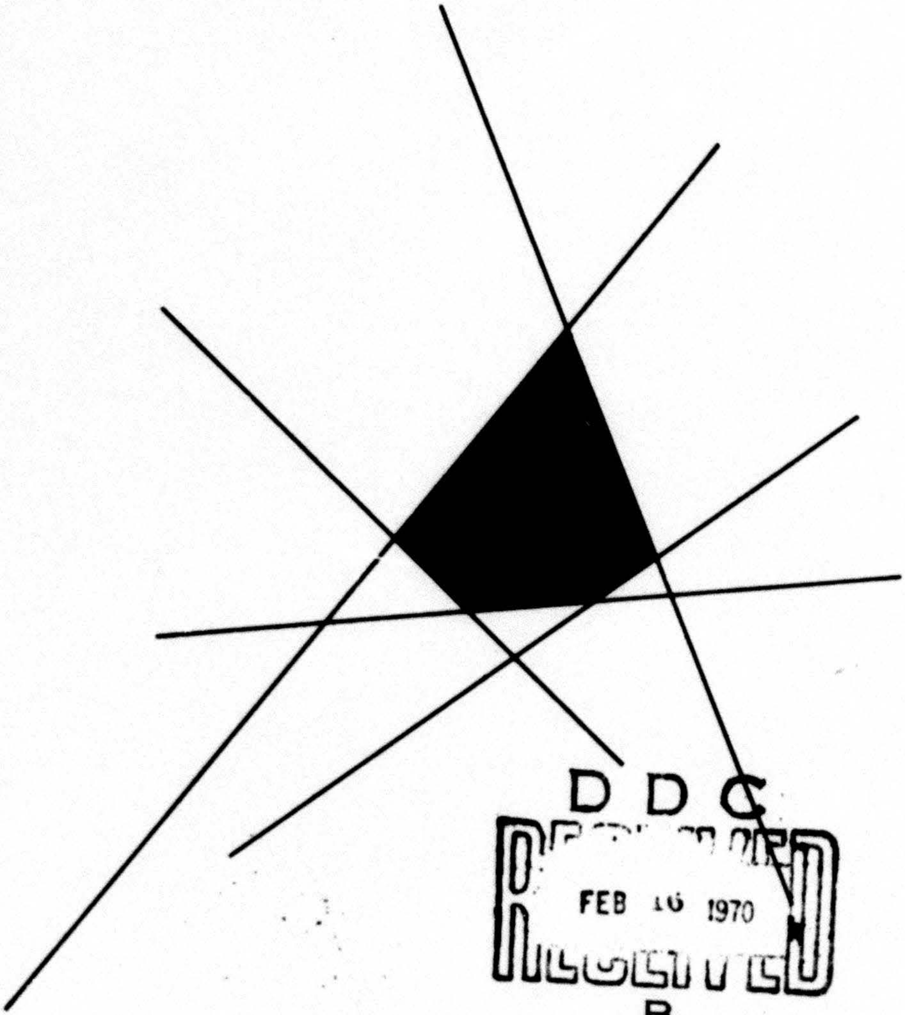
ORC 69-36
DECEMBER 1969

A NOTE ON CUTTING-PLANE METHODS WITHOUT NESTED CONSTRAINT SETS

by
DONALD M. TOPKIS



OPERATIONS RESEARCH CENTER

COLLEGE OF ENGINEERING
UNIVERSITY OF CALIFORNIA · BERKELEY

13

A NOTE ON CUTTING-PLANE METHODS WITHOUT NESTED CONSTRAINT SETS

by

Donald M. Topkis
Department of Industrial Engineering
and Operations Research
University of California, Berkeley

DECEMBER 1969 ORC 69-36

This research has been supported by the Office of Naval Research under Contract N00014-69-A-0200-1010 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

ABSTRACT

It is shown that the framework of [4] can be used to give a simplified proof of conditions given by Eaves and Zangwill [1] (which weaken the uniform concavity requirement on the objective function used by the author in [4]) under which inactive constraints may be dropped after each subproblem in cutting-plane algorithms. The convergence rate established in [4] is improved and its application extended.

A NOTE ON CUTTING-PLANE METHODS WITHOUT NESTED CONSTRAINT SETS

by

Donald M. Topkis

The problem considered is that of maximizing a real-valued continuous function f over a nonempty closed convex subset S of E^n. One is given a compact[†] convex set T containing S. The general algorithm to be considered proceeds by setting $T_o = E^n$, and, given T_k as the intersection of E^n and a finite set of closed half-spaces containing S, picking x_k to maximize f over $T_k \cap T$, stopping with x_k optimal if $x_k \in S$, and otherwise letting S_k be the intersection of E^n and a subset of the half-spaces determining T_k such that x_k maximizes f over $S_k \cap T$,[††] finding a certain closed half-space H_k containing S but not x_k, setting $T_{k+1} = S_k \cap H_k$, and continuing.

It was shown by the author in [4] that if H_k is picked in a certain manner as below and f is uniformly concave on T then $\{x_k\}$ converges to the optimum. Eaves and Zangwill [1] allowed the cuts to be certain closed convex sets (rather than just closed half-spaces) and only required (essentially) that f be quasi-concave on T and strictly quasi-concave on any convex subset of $T - S$ in proving the optimality of this procedure by using the notion of a separator. In Theorem 2 the result of Eaves and Zangwill (with the cuts here limited to certain closed half-spaces but with their weakened conditions on f) is established by a much simpler proof using the framework of [4]. Theorem 3 generalizes and extends the convergence rate established in [4].

A mapping $(a(x),b(x))$ from $T - S$ into E^{n+1} with $a(x) \in E^n$ and $b(x) \in E^1$ is a *limiting cutting-plane function* if $S \subseteq H(x) \equiv \{y : a(x)\cdot y \geq b(x)\}$

[†]This compactness assumption is relaxed in [4].

[††]If f is pseudo-concave on T then S_k would satisfy these conditions if the constraints dropped are any of those constraints determining T_k which are inactive at x_k.

for all $x \in T \sim S$, $(a(x),b(x))$ is bounded on $T \sim S$, and for any $\{x_k : k = 1,2, \ldots\} \subseteq T \sim S$ with $\lim_{k\to\infty} x_k = \bar{x} \in T \sim S$ the limit point $(\bar{a},\bar{b})$ of any convergent subsequence of $\{(a(x_k),b(x_k))\}$ satisfies $\bar{a}\cdot\bar{x} < \bar{b}$. A generalized version of this notion was introduced by Zangwill [5], and examples are given in [4] and [5].

The following was proven in [4].

__Theorem 1__:

If $H(x)$ is determined by a limiting cutting-plane function, $H_k = H(x_k)$, and $\lim_{i\to\infty} x_{k_i} = \lim_{i\to\infty} x_{k_i+1} = \bar{x}$, then $\bar{x}$ is optimal.

__Theorem 2__:

Suppose that $H(x)$ is determined by a limiting cutting-plane function, $H_k = H(x_k)$, and f is quasi-concave on T and strictly quasi-concave on any convex subset of $T \sim S$. Then the limit point of any convergent subsequence of $\{x_k\}$ is optimal.

__Proof__:

Let $\bar{x}$ be the limit point of any convergent subsequence of $\{x_k\}$. Since $\{x_k\}$ is bounded there then exists a subsequence $\left\{x_{k_i}\right\}$ such that $\lim_{i\to\infty} x_{k_i} = \bar{x}$ and $\lim_{i\to\infty} x_{k_i+1} = \bar{y}$. Now suppose $\bar{x}$ is not optimal. If $\bar{x}$ were feasible it would be optimal [4] so $\bar{x} \notin S$, and by Theorem 1, $\bar{x} \neq \bar{y}$. Since x_{k_i} maximizes f over the convex set $S_{k_i} \cap T$ and $x_{k_i+1} \in T_{k_i+1} \cap T \subseteq S_{k_i} \cap T$ for all i , $f\left(x_{k_i}\right) \geq f\left(\alpha x_{k_i} + (1-\alpha)x_{k_i+1}\right)$ for all i and all $\alpha \in [0,1]$. By continuity,

$$f(\bar{x}) \geq f(\alpha\bar{x} + (1-\alpha)\bar{y}) \qquad \text{for all } \alpha \in [0,1] . \tag{1}$$

Since $f(x_k)$ is nonincreasing in k it is easily seen that $f(\bar{x}) = f(\bar{y})$ so by quasi-concavity

$$f(\alpha\bar{x} + (1-\alpha)\bar{y}) \geq \min\{f(\bar{x}), f(\bar{y})\} = f(\bar{x}) \quad \text{for all } \alpha \in [0,1]\ . \tag{2}$$

By (1) and (2),

$$f(\bar{x}) = f(\alpha\bar{x} + (1-\alpha)\bar{y}) \quad \text{for all } \alpha \in [0,1]\ . \tag{3}$$

Since S is closed and $\bar{x} \notin S$, there exists $\gamma \in [0,1)$ such that $\alpha\bar{x} + (1-\alpha)\bar{y} \notin S$ for all $\alpha \in [\gamma,1]$. But by (3) and the strict quasi-concavity of f on the line segment joining $\gamma\bar{x} + (1-\gamma)\bar{y}$ and $\bar{x}$,

$$f(\alpha\bar{x} + (1-\alpha)\bar{y}) > \min\{f(\bar{x}), f(\gamma\bar{x} + (1-\gamma)\bar{y})\} = f(\bar{x}) \quad \text{for all } \alpha \in (\gamma,1)\ . \tag{4}$$

But (4) contradicts (3), so $\bar{x}$ must be optimal. ||

Levitin and Polyak [3] have established an arithmetic convergence rate for a cutting-plane algorithm which, when specialized to subsets of E^n, has $S_k = T_k$ (although their proof still holds if inactive constraints were dropped after each subproblem) and uses the cutting-plane method of Lemma 3 of [4] (i.e., $\alpha(x) = 1$ always, below). Here their algorithm and method of proof are generalized to show the same convergence rate for algorithms which allow inactive constraints to be dropped after each subproblem and for which the cutting plane at $x \in T - S$ may be generated at some point $w(x) \in \overline{T - S}$ other than x on the line segment joining x to a certain interior point of S (as in Lemma 4 of [4]). The following eliminates the restriction of the generalization given in [4] that $G(w(x)) \leq \varepsilon G(x)$ for some $\varepsilon \in (0,1]$ and the dependence of the convergence rate on ε.

Theorem 3:

Suppose that $S = \{x : G(x) \geq 0,\ x \in T\}$, $G(x)$ is concave and continuous on T, there exists $t \in S$ with $G(t) > 0$, and for $x \in T - S$ define $\lambda(x) = \sup\{\lambda : \lambda x + (1-\lambda)t \in S\}$ and set $w(x) = \alpha(x)x + (1-\alpha(x))t$ for any $\alpha(x) \in [\lambda(x),1]$. Suppose also that there exists a function $\mu(w)$ from $\overline{T - S}$

into E^n with $|\mu(w)| \le K$ for all $w \in \overline{T \sim S}$ and such that $G(y) \le G(w) + \mu(w)\cdot(y - w)$ for all $w \in \overline{T \sim S}$ and $y \in S$, and let $H_k = \{x : 0 \le G(w(x_k)) + \mu(w(x_k))\cdot(x - w(x_k))\}$. If f is strongly concave on T and differentiable on S and $\bar{x}$ is the unique maximum of f on S, then for $k \ge 1$

$$f(x_k) - f(\bar{x}) \le \frac{1}{a_1 k}$$

and

$$|x_k - \bar{x}| \le \frac{1}{a_2\sqrt{k}}$$

where

$$a_1 = 2\gamma\left(\frac{G(t)}{bdK}\right)^2 , \quad a_2 = \frac{2\gamma G(t)}{bdK} ,$$

$d = \max\{|\nabla f(y)| : y \in S\}$, $b = \max\{|y - t| : y \in T\}$, and γ is as in the definition of strong concavity.

<u>Proof</u>:

Let $\lambda_k = \lambda(x_k)$, $\alpha_k = \alpha(x_k)$, $w_k = w(x_k)$, and $\mu_k = \mu(w_k)$.

Clearly $\lambda_k x_k + (1 - \lambda_k)t = \frac{\lambda_k}{\alpha_k} w_k + \left(1 - \frac{\lambda_k}{\alpha_k}\right)t$ and $G(\lambda_k x_k + (1 - \lambda_k)t) = 0$, so by concavity

$$0 = G\left(\frac{\lambda_k}{\alpha_k} w_k + \left(1 - \frac{\lambda_k}{\alpha_k}\right)t\right) \ge \frac{\lambda_k}{\alpha_k} G(w_k) + \left(1 - \frac{\lambda_k}{\alpha_k}\right)G(t)$$

and

$$\left(1 - \frac{\alpha_k}{\lambda_k}\right)G(t) \ge G(w_k) . \tag{5}$$

Since $t \in S$,

$$G(t) \leq G(w_k) + \mu_k \cdot (t - w_k) . \tag{6}$$

But it is easily seen that $x_k \notin H_k$ so $x_k \neq x_{k+1} \in T_{k+1} \cap T = H_k \cap S_k \cap T$ and by the strict concavity of f on T ,

$$0 = G(w_k) + \mu_k \cdot (x_{k+1} - w_k) \tag{7}$$

or

$$0 = G(w_k) + \alpha_k \mu_k \cdot (x_{k+1} - x_k) + (1 - \alpha_k)\mu_k \cdot (x_{k+1} - t) . \tag{8}$$

By (6) and (7)

$$-G(t) \geq \mu_k \cdot (x_{k+1} - t) \tag{9}$$

so by (8), (9), and (5)

$$\begin{aligned} 0 &\leq G(w_k) + \alpha_k \mu_k \cdot (x_{k+1} - x_k) - (1 - \alpha_k)G(t) \\ &\leq \left(1 - \frac{\alpha_k}{\lambda_k}\right)G(t) + \alpha_k K|x_{k+1} - x_k| - (1 - \alpha_k)G(t) \\ &= -\frac{\alpha_k}{\lambda_k}(1 - \lambda_k)G(t) + \alpha_k K|x_{k+1} - x_k| \\ &\leq -\alpha_k(1 - \lambda_k)G(t) + \alpha_k K|x_{k+1} - x_k| \end{aligned}$$

or

$$1 - \lambda_k \leq \frac{K}{G(t)} |x_{k+1} - x_k| . \tag{10}$$

By the concavity of f and the optimality of $\bar{x}$ on S ,

$$(11)\qquad f(x_k) - f(\bar{x}) \le f(x_k) - f(\lambda_k x_k + (1-\lambda_k)t)$$
$$\le (1-\lambda_k)(x_k - t)\cdot\nabla f(\lambda_k x_k + (1-\lambda_k)t)$$
$$\le (1-\lambda_k)|x_k - t|\cdot|\nabla f(\lambda_k x_k + (1-\lambda_k)t)|$$
$$\le (1-\lambda_k)bd\ .$$

Combining (10) and (11),

$$(12)\qquad f(x_k) - f(\bar{x}) \le \frac{bdK}{G(t)}|x_{k+1} - x_k|\ .$$

Since $x_k, x_{k+1} \in S_k \cap T$ and x_k maximizes the strongly concave function f over the convex set $S_k \cap T$, for some $\gamma > 0$

$$(13)\qquad f(x_k) > f(\tfrac{1}{2}(x_k + x_{k+1})) \ge \tfrac{1}{2}f(x_k) + \tfrac{1}{2}f(x_{k+1}) + \gamma|x_k - x_{k+1}|^2$$

and from (13)

$$(14)\qquad f(x_k) - f(x_{k+1}) \ge 2\gamma|x_k - x_{k+1}|^2\ .$$

Now let $D_k = f(x_k) - f(\bar{x}) > 0$. By (12) and (14)

$$D_k^2 \le \left(\frac{bdK}{G(t)}\right)^2|x_{k+1} - x_k|^2 \le \left(\frac{bdK}{G(t)}\right)^2\left(\frac{1}{2\gamma}\right)(D_k - D_{k+1})$$

or

$$(15)\qquad D_{k+1} \le D_k - a_1 D_k^2 = D_k(1 - a_1 D_k)\ .$$

The arithmetic convergence rate for D_k then follows from (15) as in [2] by observing that

(16) $$\frac{1}{D_{k+1}} \geq \frac{1}{D_k}\left(\frac{1}{1 - a_1 D_k}\right) = \frac{1}{D_k}\left(\sum_{i=0}^{\infty} (a_1 D_k)^i\right)$$

$$\geq \frac{1}{D_k}(1 + a_1 D_k) = \frac{1}{D_k} + a_1$$

and using induction on (16) to get

$$\frac{1}{D_k} \geq \frac{1}{D_o} + a_1 k$$

or

(17) $$D_k \leq \frac{1}{\frac{1}{D_o} + a_1 k} \leq \frac{1}{a_1 k} .$$

As in (13) and (14), it follows that

(18) $$D_k = f(x_k) - f(\bar{x}) \geq 2\gamma |x_k - \bar{x}|^2 ,$$

and by (17) and (18)

$$|x_k - \bar{x}| \leq \frac{1}{\sqrt{2\gamma a_1 k}} . \quad ||$$

REFERENCES

[1] Eaves, B. C. and W. I. Zangwill, "Generalized Cutting Plane Algorithms," Working Paper No. 274, Center for Research in Management Science, University of California, Berkeley, (July 1969). (To appear in SIAM Journal on Control.)

[2] Frank, M. and P. Wolfe, "An Algorithm for Quadratic Programming," Naval Research Logistics Quarterly, Vol. 3, pp. 95-110, (1956).

[3] Levitin, E. S. and B. T. Polyak, "Constrained Minimization Methods," Zh.vychisl.Mat.mat.Fiz. (in Russian), Vol. 6, pp. 787-823, (1966); also, U.S.S.R. Computational Mathematics and Mathematical Physics (in English), pp. 1-50, (1968).

[4] Topkis, D. M., "Cutting-Plane Methods Without Nested Constraint Sets," ORC 69-14, Operations Research Center, University of California, Berkeley, (June 1969). (To appear in Operations Research.)

[5] Zangwill, W. I., NONLINEAR PROGRAMMING: A UNIFIED APPROACH, Prentice-Hall, Inc., (1969).

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of California, Berkeley	2a. REPORT SECURITY CLASSIFICATION Unclassified
	2b. GROUP

3. REPORT TITLE

A NOTE ON CUTTING-PLANE METHODS WITHOUT NESTED CONSTRAINT SETS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Research Report

5. AUTHOR(S) (First name, middle initial, last name)

Donald M. Topkis

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
December 1969	8	5

8a. CONTRACT OR GRANT NO. N00014-69-A-0200-1010	9a. ORIGINATOR'S REPORT NUMBER(S) ORC 69-36
b. PROJECT NO. NR 047 033	
c. Research Project No.: RR 003 07 01	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
NONE	Mathematical Science Division, Office of Naval Research Washington, D.C. 20360

13. ABSTRACT

SEE ABSTRACT.

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6811

Unclassified
Security Classification

A-31408

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Nonlinear Programming						
Algorithms						
Cutting-Plane Algorithms						
Convergence Rates						

DD FORM 1 NOV 65 1473 (BACK)

S/N 0101-807-6821

Unclassified

Security Classification

A-31409